with user **102**'s family members or coworkers). In step **704**, audio samples are received from the identified devices. For example, audio samples **400**, **402**, and **404** may respectively be received from secondary devices **106**, **108**, and **110**. In step **706**, the received audio samples may be compared to a reference sample of user **102**'s voice to identify samples or portions of samples that contain voices other than user **102**'s voice, and the extraneous samples (or extraneous portions of the samples) may be discarded. In step **708**, an audio sample may be selected from among the audio samples based on its suitability for speech recognition. For example, multi-device speech recognition apparatus **200** may select audio sample **400**, from among audio samples **400**, **402**, and **404**, based on its suitability for speech recognition.

[0036] The methods and features recited herein may be implemented through any number of computer readable media that are able to store computer readable instructions. Examples of computer readable media that may be used include RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, DVD or other optical disk storage, magnetic cassettes, magnetic tape, magnetic storage and the like.

[0037] Additionally or alternatively, in at least some embodiments, the methods and features recited herein may be implemented through one or more integrated circuits (ICs). An integrated circuit may, for example, be a microprocessor that accesses programming instructions or other data stored in a read only memory (ROM). In some embodiments, a ROM may store program instructions that cause an IC to perform operations according to one or more of the methods described herein. In some embodiments, one or more of the methods described herein may be hardwired into an IC. In other words, an IC may comprise an application specific integrated circuit (ASIC) having gates and other logic dedicated to the calculations and other operations described herein. In still other embodiments, an IC may perform some operations based on execution of programming instructions read from ROM or RAM, with other operations hardwired into gates or other logic. Further, an IC may be configured to output image data to a display buffer.

[0038] Although specific examples of carrying out the disclosure have been described, those skilled in the art will appreciate that there are numerous variations and permutations of the above-described apparatuses and methods that are contained within the spirit and scope of the disclosure as set forth in the appended claims. Additionally, numerous other embodiments, modifications, and variations within the scope and spirit of the appended claims may occur to persons of ordinary skill in the art from a review of this disclosure. Specifically, any of the features described herein may be combined with any or all of the other features described herein.

1-35. (canceled)

36. A method comprising:

identifying one or more secondary devices in physical proximity to a user of a principal device, each of the one or more secondary devices being configured to capture audio;

receiving a plurality of audio samples captured by the one or more secondary devices; and

selecting an audio sample comprising a voice of the user of the principal device from among the plurality of audio samples captured by the one or more secondary devices based on suitability of the audio sample for speech recognition.

37. The method of claim **36**, wherein identifying the one or more secondary devices in physical proximity to the user of the principal device comprises:

receiving current location information from each of a predetermined set of secondary devices; and

identifying the one or more secondary devices in physical proximity to the user of the principal device by comparing the current location information received from each of the predetermined set of secondary devices with current location information for the principal device to determine which of the predetermined set of secondary devices are physically proximate to the principal device.

38. The method of claim **36**, wherein selecting the audio sample comprising the voice of the user of the principal device comprises:

converting, via speech recognition, the plurality of audio samples into a plurality of corresponding text strings;

determining a plurality of recognition confidence values, each of the plurality of recognition confidence values corresponding to a level of confidence that a corresponding text string of the plurality of corresponding text strings accurately reflects content of an audio sample of the plurality of audio samples from which the corresponding text string was converted;

identifying, from among the plurality of recognition confidence values, a recognition confidence value indicating a level of confidence as great or greater than that of each of the plurality of recognition confidence values; and

selecting an audio sample of the plurality of audio samples that corresponds to the identified recognition confidence value indicating the level of confidence as great or greater than that of each of the plurality of recognition confidence values.

39. The method of claim **36**, wherein selecting the audio sample comprising the voice of the user of the principal device during the period of time comprises:

analyzing the plurality of audio samples to identify an audio sample of the plurality of audio samples that is equally well suited or more well suited for speech recognition; and

selecting the identified audio sample of the plurality of audio samples that is equally well suited or more well suited for speech recognition.

40. The method of claim **39**, wherein analyzing the plurality of audio samples to identify the audio sample of the plurality of audio samples that is equally well suited or more well suited for speech recognition comprises at least one of:

determining a plurality of signal-to-noise ratios, each of the plurality of signal-to-noise ratios corresponding to one of the plurality of audio samples, and wherein the audio sample of the plurality of audio samples that is equally well suited or more well suited for speech recognition corresponds to a signal-to-noise ratio of the plurality of signal-to-noise ratios that indicates a proportion of signal-to-noise that is as great or greater than each of the plurality of signal-to-noise ratios;

determining a plurality of amplitude levels, each of the plurality of amplitude levels corresponding to one of the plurality of audio samples, and wherein the audio sample of the plurality of audio samples that is equally